DOCUMENT RESUME

ED 415 206 SP 037 715

AUTHOR Ingersoll, Richard M.
TITLE National Assessments of Teacher Quality. Working Paper Series, Working Paper No. 96-24.
INSTITUTION American Institutes for Research in the Behavioral Sciences, Washington, DC.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-WP-96-24
PUB DATE 1996-10-00
NOTE 42p.
AVAILABLE FROM U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Avenue, N.W.; Room 400, Washington, DC 20208-5654.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Quality; Elementary School Teachers; Elementary Secondary Education; *Evaluation Methods; Knowledge Base for Teaching; *National Surveys; Peer Evaluation; Secondary School Teachers; Self Evaluation (Individuals); *Social Influences; Sociology; Teacher Competencies; *Teacher Effectiveness; *Teacher Evaluation
IDENTIFIERS *Alternative Assessment; National Teacher Examinations; Professionalization of Teaching

ABSTRACT

This report discusses evaluation and measurement of the quality of school teachers, emphasizing the problem of assessing competence, performance, and effectiveness through large-scale national sample surveys. Part 1 outlines types of issues and questions surrounding teacher quality that must be addressed, highlighting conventional and new approaches to teacher assessment. Part 2 provides an interpretive and critical review of the predominant approaches to assessing competence, performance, and effectiveness of teachers. Part 3 offers an alternative, sociological approach to assessment. This view presents employee and organizational assessment as social phenomena, suggesting that all evaluation is shaped by social context, is value-laden and purposeful, and is based on particular norms of appropriate behavior. In this view, the role of research should be to uncover viewpoints, interests, criteria, and values underlying and shaping any particular theory and assessment method in order to illuminate its strengths and limits. The goal of research should be to help develop a range of complementary methods representing a range of viewpoints. Part 4 draws out implications of these approaches for large-scale national survey assessments of teacher quality. Part 5 lists sources and examples of existing survey measures of teacher quality (self-assessment, and administrator and peer assessment). (Contains 58 references.) (SM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 415 206

National Center for Education Statistics

National Assessments of Teacher Quality

Working Paper No. 96-24

October 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

U.S. Department of Education
Office of Educational Research and Improvement

BEST COPY AVAILABLE

National Assessments of Teacher Quality

Working Paper No. 96-24 October 1996

Contact: Dan Kasprzyk
Surveys and Cooperative Systems Group
(202) 219-1588
e-mail: daniel_kasprzyk@ed.gov

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Pascal D. Forgione, Jr.
Commissioner

Surveys and Cooperative Systems Group
Paul D. Planchon
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *National Assessments of Teacher Quality*, Working Paper No. 96-24, by Richard M. Ingersoll. Project Officer, Daniel Kaspryzk. Washington, D.C.: 1996.

October 1996

Foreword

Each year a large number of written documents are generated by NCES staff and individuals commissioned by NCES which provide preliminary analyses of survey results and address technical, methodological, and evaluation issues. Even though they are not formally published, these documents reflect a tremendous amount of unique expertise, knowledge, and experience.

The ***Working Paper Series*** was created in order to preserve the information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series. Consequently, we encourage users of the series to consult the individual authors for citations.

To receive information about submitting manuscripts or obtaining copies of the series, please contact Ruth R. Harris at (202) 219-1831 or U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave., N.W., Room 400, Washington, D.C. 20208-5654.

Susan Ahmed
Chief Mathematical Statistician
Statistical Standards and
Services Group

Samuel S. Peng
Director
Methodology, Training, and Customer
Service Program

National Assessments of Teacher Quality

Richard M. Ingersoll

October 1996

Department of Sociology
University of Georgia
and
American Institutes for Research
Washington, D.C.

Table of Contents

Page

FOREWORD iii

ACKNOWLEDGMENTS vi

INTRODUCTION 1

BACKGROUND 2

APPROACHES TO TEACHER ASSESSMENT 4
- Conventional Approaches to Teacher Assessment 4
- New Approaches to Teacher Assessment 7

A SOCIOLOGICAL APPROACH TO TEACHER ASSESSMENT 16

IMPLICATIONS FOR NATIONAL ASSESSMENTS 20

SOURCES OF SURVEY MEASURES OF TEACHER QUALITY 22
- Self Assessment 22
- Administrator and Peer Assessment 23

CONCLUSION 25

REFERENCES 26

ACKNOWLEDGMENTS

This paper greatly benefitted from the suggestions of Rick White of AIR, Mary Kennedy of Michigan State University, Hunter Moorman of the Office of Research of OERI, and Marilyn McMillen, John Burkett, Sharon Bobbitt, Shelley Burns, Saul Carrallo and Dan Kasprzyk of NCES. Partial support for this project was provided by a Research Fellowship jointly sponsored by the National Science Foundation and the National Center for Education Statistics and administered by the American Educational Research Association.

INTRODUCTION

This report addresses the issue of evaluating and measuring the quality of school teachers. In particular, it speaks to the problem of assessing the competence, performance and effectiveness of elementary and secondary teachers through large-scale national sample surveys. Its objective is to provide a foundation and springboard for thinking about the conceptual and methodological issues underlying the development of national survey measures of teacher quality. This paper will not attempt to conceive or construct such measures or indicators themselves. Rather, it seeks to provide background to such efforts. Its role is to review the range of contemporary thought on assessing teacher quality and to discuss the possibilities and problems of adapting existing methods and measures for use in large-scale surveys, such as those undertaken by the National Center for Education Statistics (NCES) of the Office of Educational Research and Improvement (OERI) of the U.S. Department of Education.

This report begins by outlining in more detail the kinds of issues and questions surrounding teacher quality that need to be addressed. Part 2 then provides an interpretive and critical review of the predominant approaches to assessing the competence, performance and effectiveness of teachers. Part 3 offers an alternative - a sociological approach to assessment. In this view, employee and organizational assessment are social phenomena. That is, all evaluation is shaped by its social context, is value-laden, purposeful and is based on particular norms of appropriate behavior. The role of research, I argue, should be to uncover the viewpoints, interests, criteria, and values underlying and shaping any particular theory and method of assessment, in order to illuminate its strengths and limits. Rather than seek "the one best way," or to strive to develop "objective" or "authentic" methods of assessment, the goal of research should be, I suggest, to help develop a range of complementary methods, representing a range of viewpoints. Part 4 draws out some implications of these approaches for large-scale national survey assessments of teacher quality. Finally, part 5 concludes this report by listing some sources and examples of existing survey measures of teacher quality.

BACKGROUND

This paper addresses a larger effort by NCES, OERI and the Department of Education as a whole, to review, evaluate and rethink the national education data and information system. Two overview reports - A Guide to Improving the National Education Data System and Education Counts have been devoted to this larger end. The first was produced in 1990 by the National Education Statistics Agenda Committee (NESAC) of the National Forum on Education Statistics - an organization, created in 1989, of a range of both government agencies and independent associations responsible for collecting, reporting and using national educational information. The second was produced in 1991 by the Special Study Panel on Education Indicators - a second group, chartered in 1989 by the Department of Education.

Both the Panel and the Forum were partly developed in response to the Hawkins-Stafford Elementary and Secondary School Improvement Amendments passed by Congress in 1988 and in response to the context of enhanced national attention to school reform, exemplified by the National Education Goals effort initiated by President Bush and the nation's governors in 1989. Both the Panel and the Forum were authorized to examine the overall purposes behind national education data collection, in general, and the role of NCES, as the key national education data collection agency, in particular.

Both groups' reports specifically reviewed the kinds of national education data needed, the kinds of data currently available and the gaps between the two. In respect to the latter, both reports stressed the need to gather more and better data and information on the quality of elementary and secondary schools in the U.S.

There has been, of course, extensive national data collected on many key indicators of school quality, such as expenditures per pupil, the number of books in libraries, the availability of science laboratories, and the range of services for special populations. The reports agreed these measures are important, but both also agreed that alone these are not adequate measures of the quality and character of schools. A key gap in the national data system, the reports concluded, is the dearth of information on the actual processes and practices within schools.

The paucity of this latter type of school quality data stems from two sources. First, there has been little consensus about both how to define and how to accurately measure this domain. Second, there is little precedent for the collection of such data at the national level. Having said that, both reports have recommended efforts be made to collect far more and far better information on this domain.

More specifically, within the domain of school quality, both reports drew attention to the need to define and to measure more accurately the quality of teachers. The reports provided evidence of the growing consensus that assessment of this subdomain presents both one of the most important and also one of the most intractable problems in the development of education indicators.

Teacher quality is a complex phenomenon. It comprises at least two distinct elements: 1.) teacher qualifications and 2.) teaching quality. The first refers to the innate competencies teacher candidates bring to the job and the kinds, amounts and caliber of training these candidates do or do not receive prior to or during their career. The second refers to the actual caliber of the teaching the teacher does, once on the job.

As indicated by a follow-up and third report specifically detailing the most significant gaps in the national data system, OERI and NCES currently collect extensive data on indicators associated with the first element of teacher quality, including the academic performance, background, preparation, training, and certification of both new and experienced teachers (Bobbitt, Quinn and Dabbs 1992). But, the greatest difficulty and most substantial need lies with measuring the second element - the actual practices and behavior of teachers in classrooms. Again, largely because there is little consensus on what it is and how best to measure it, there has been ambivalence about efforts to collect national data on measures of teacher practices such as, the extent to which teachers employ effective instructional and classroom techniques, and the degree of commitment and responsibility of teachers to their work.

Although data on teacher practices are both scarce and difficult to obtain, this is a variable of crucial importance because it may be the most important aspect of teacher quality. Whether a teacher has high innate abilities or excellent training, ultimately, may be less important than how well teachers actually perform in classrooms. In turn, teacher quality may be one of the most important factors shaping the overall quality and performance of schools. And, ultimately, school quality is undoubtedly a key factor in how well students learn, grow and perform.

The above mentioned reports drew attention to both the lack of and the importance of data on teacher quality. The Guide, for instance, recommended that OERI fund special studies specifically to improve the measurement of teacher competence and efforts have been initiated at OERI and NCES to address these ongoing needs.

This paper represents one such effort. It is designed to aid the background thinking necessary to understanding two key questions: 1.) how to define and, 2.) how to measure the quality of teachers. Below, I begin by critically reviewing how educational researchers and school officials have currently been attempting to answer these two questions.

APPROACHES TO TEACHER ASSESSMENT

Assessment and evaluation of how well elementary and secondary school teachers teach has been a recurrent concern since the initial development of the nation's educational system in the nineteenth century. School officials, education policy makers, researchers and parents have all had a great deal of interest in both gauging and improving the quality of teachers and teaching. This is not surprising. Elementary and secondary schooling are mandatory in the U.S. and it is into the custody of teachers that children are legally placed for a significant portion of their lives. Moreover, the quality of teachers and teaching are undoubtedly among the most important factors shaping the overall achievement and growth of students.

This concern with the quality of teachers and schools, however, has seen a dramatic rise in the past two decades. Beginning in the 1970's, the number and variety of methods to assess and evaluate teachers - their abilities, preparation, training, and performance - has greatly expanded (Sclan 1994). Most of these efforts have been conducted at the state, district or school levels. In most cases, their purposes have been quite different than that behind the newer interest in national survey assessments, such as that contemplated by OERI and NCES. Rather than designed to provide an overall portrait of the caliber of teachers across the nation, current teacher assessment is largely devoted to highly applied ends - teacher entry, licensure, certification, promotion, tenure, merit pay awards, termination, and staff development - and in specific settings.

As teacher assessment has increased in importance, it has, however, become more controversial. Indeed, research, policy and practice concerned with teacher assessment is marked by a great deal of disagreement. This disagreement largely surrounds two key questions underlying the assessment and evaluation of teacher quality - what is to be measured and how best to do it (Haertel 1991; Haney et al. 1987; Millman and Darling-Hammond 1990).

As discussed below, different approaches to teacher assessment hold very different conceptions of what the process of teaching actually involves, and hence, what are the key characteristics of the good or effective teacher. Moreover, different approaches turn to different methods for how to best measure these key characteristics.[1]

Conventional Approaches to Teacher Assessment

Until recently, the predominant approach to teacher assessment has viewed such evaluation as an issue of employee accountability. A key factor driving this approach is the public perception that school problems are, to an important extent, teacher problems - that is, there are significant inadequacies in the ability, training, motivation and performance of teachers in the U.S.

---

[1]This paper has drawn on a number of good in-depth overviews of research and policy concerned with teacher assessment. Among those I found most useful were: Dwyer and Stufflebeam 1992; Haertel 1991; Haney et al. 1987 and Kennedy 1992. In addition, for an excellent collection of articles see the New Handbook of Teacher Evaluation, edited by Millman and Darling-Hammond (1990).

Moreover, there is a widespread perception that schools either cannot or will not correct these inadequacies. In particular, schools do not seem to "weed out" incompetent teachers. The result, over the past two decades, has been a growing demand for and large growth in the use of teacher assessment to enhance the accountability of teachers as public employees (Educational Research Service 1988; Dwyer and Stufflebeam 1992; Kirst 1990). Several methods have been used.

The first and perhaps the most traditional method of teacher assessment is classroom observation of individual teachers, usually conducted by school administrators or supervisors. These are usually referred to as classroom performance assessments. In this method, an evaluator typically spends several class periods observing the teacher at work and grades him or her by utilizing a standard checklist of appropriate teacher practices.

A second method of teacher assessment is the use of written tests or examinations administered to teachers themselves. Unlike classroom observations, these pencil and paper tests do not directly assess teaching performance. Rather, they are designed to measure a teacher's basic literacy and numeracy skills, and subject matter knowledge in particular areas. The most common is the National Teacher Examination (NTE), produced by the Educational Testing Service (ETS). Their overall use has dramatically increased, as of the late 1980s, more than half of the states used all or part of the National Teacher Examination in teacher assessment (OERI 1987).

A final method uses student performance to assess teacher performance. In this case, a teacher's performance is judged by gains in their students' academic achievement, as measured on standardized achievement tests. These have been used to compare the effectiveness of teachers within or between schools or school districts.

In theory, these methods of evaluation are designed to insure that both the qualifications and performance of teachers are at an adequate level and also to instill a general sense of accountability in the teaching workforce, and hence improve teacher quality. As a result, these methods of teacher assessment have gained in popularity and both policy makers and education officials have increasingly instituted their use at the school, district and state levels (Educational Research Service 1988; Dwyer and Stufflebeam 1992). However, despite this widespread acceptance, there has been little, if any, evidence that these testing and classroom observation methods have improved the quality of the teaching force (Haertel 1991; Haney et al. 1987). In fact, all have come under criticism from a number of quarters. Critics have taken issue with both the theory and methods of such programs. Moreover, a number of court and legal challenges to the equity and accuracy of these assessment methods have clouded the legality of school officials' use of them for teachers' employment and promotion decisions (Haertel 1991).

One set of criticisms surrounds the conception and definition of the teaching processes underlying these methods of assessment. All the above-described methods - classroom observations, teacher examinations, and student performance measures - have been criticized for subscribing to both a narrow and a shallow view of what the work of teaching entails and, hence,

what constitutes effective teaching.

Teacher exams, on the one hand, focus on the "what" of teaching - academic subject knowledge. They usually include only a small number of items devoted to the "how" of teaching - pedagogical knowledge and skills. While most agree that having basic subject knowledge is an important prerequisite to effective teaching, critics have argued this is certainly not a sufficient indication of the range of knowledge and skills needed to instruct and manage groups of children (e.g. Darling-Hammond 1986). Hence, many have concluded that teacher exams do not actually measure a teacher's ability to teach (Madaus and Pullin 1987).

The checklists, commonly used in classroom performance assessments, on the other hand, focus almost exclusively on pedagogical skills, as opposed to subject knowledge. These instruments are designed to measure practices and attitudes thought to be associated with effective teaching, such as, eye contact, enthusiasm, time on task, and avoidance of negative reinforcement. But, in this case, critics have argued that many of the variables measured on checklists are trivial and superficial. They hold that such checklists do not capture many of the most crucial and sophisticated aspects of teaching, such as ability to interact with parents, test construction, grading criteria, lesson planning, managing classrooms, ability to communicate, and knowledge of the needs and capacities of different age levels of children. The result, according to the critics, is that classroom performance assessments often focus on teaching style, rather than substance (Shulman 1986; Haertel 1991; Stodolsky 1984; Murnane and Raizen 1988).

Moreover, critics have held that in classroom observations, school administrators typically utilize standardized pre-made observation forms that, in effect, allow evaluators to bypass the time consuming, but all important, preliminary task of clarifying what are effective teaching practices in their schools. Critics term this the "law of the instrument" - the criteria of effective teaching are, by default, those underlying the most convenient and available measurement instruments (Dwyer and Stufflebeam 1992).

The use of student achievement test score gains to assess teachers has also been criticized for the conception of teaching and learning such tests assume. Standardized student achievement tests assess minimum levels of student competence, overlook non-academic aspects of student learning and are limited to the kinds of knowledge that can be captured with multiple-choice formats. Critics have pointed out that effective teaching includes a far wider range of skills than simply teaching what is measured on such tests (Dwyer and Stufflebeam 1992; Scriven 1991; Murnane and Raizen 1988).

Moreover, all three assessment methods have also been criticized for underestimating differences in the knowledge and skills required to teach different subjects and grade levels. For instance, the performance checklists, typically used in observational methods, assume a core of knowledge generic to all teachers, to all subjects and to all grade levels. These overlook, critics claim, the enormous differences in the knowledge of students, of subject matter and of teaching methods required by teachers at different grade levels (Madus and Pullin 1987; Stodolsky 1984;

Darling-Hammond 1986; Shulman 1986).

Along with the breadth and depth of the conception of teaching underlying conventional forms of assessment, a second set of criticisms surrounds the quality and accuracy of the methods themselves. Numerous analysts have argued that conventional assessment methods suffer from serious problems of accuracy.

For instance, the use of student achievement test score gains to assess teachers has been severely criticized for the inability to separate out the portion of student achievement gains that are actually attributable to specific teachers. There are numerous other factors that could also effect student achievement, such as home background, student personality, attendance, school resources, the peer group, community attitudes and the socio-economic status of the students' families. Assessments that do not control for all these other potential factors may hold teachers accountable for things they are unable to influence and, hence, for results not of their own making (Millman 1981; Berk 1988; Dwyer and Stufflebeam 1992).

In addition, school administrators charged with evaluating teachers with classroom performance checklists often have no training in evaluation, may know little of the particular subject being taught, and may face a natural conflict of interest between finding fault with a teacher and developing communication with a future colleague. Possibly for these reasons, teachers' performance assessments have been found to lack variability; many administrators simply give most teachers good evaluations (Medley and Coker 1987; Choy et al. 1993, p. 104).

In sum, as these methods of teacher assessment have become more popular in recent years, they have been subject to an array of serious criticisms on both conceptual and methodological grounds. Critics assert that the most common teacher assessment methods are based on overly simplistic prescriptions for effective teaching; that is, they focus on knowledge and skills that may not be necessary for effective teaching and they omit many of the critical and the most important aspects of teachers' work. Moreover, critics have also charged that many of these instruments do not produce accurate measures; that is, they do not measure what they are supposed to measure with an adequate degree of consistency.

New Approaches to Teacher Assessment

Although the above criticisms of conventional teacher assessment methods take a number of forms and come from a number of different quarters, there is a common theme running through much of the debate. Underlying the resistance to the conventional modes of teacher assessment is the notion that the road to improvements in teacher quality will not come through increasing the scrutiny and accountability of teachers. There is a growing consensus among educators, researchers and policy makers, that if teaching is to be improved, an entirely different approach to assessment must be developed. In this view, rather than subjecting teachers to greater control, scrutiny and accountability, the objective of assessment should be to foster the ongoing personal

and professional growth and development of teachers. Moreover, in this view, rather than something imposed on teachers, assessment must be something in which teachers have a hand in creating, administering and using.

This newer view of teacher assessment is bound up with a larger movement in the realm of education reform that has dramatically grown since the mid 1980's - teacher professionalization. There has been a growing consensus among education reformers, policymakers and researchers that many of the well-publicized shortcomings of the elementary and secondary education system in the U.S. are, to an important extent, due to inadequacies in the working conditions, resources and support afforded to school teachers. Proponents of this view have argued, for example, that teachers are underpaid, have too little say in the operation of schools, have too few opportunities to improve their teaching skills, suffer from a lack of support or assistance and are not adequately rewarded or recognized for their efforts. The key to improving the quality of schools, these critics hold, lies in upgrading the status, training and working conditions of teaching, that is, in furthering the professionalization of teachers and teaching. The rationale underlying this view is that upgrading the teaching occupation will lead to improvements in the motivation and efficacy of teachers, which, in turn, will lead to improvements in teachers' performance, which will ultimately lead to improvements in student learning.[2]

One of the primary targets of the teacher professionalization movement has been the need for new forms of teacher assessment. In this view, assessment must be built on a more sophisticated conception of what the work of teachers entails and what constitutes effective teaching. In turn, more authentic methods of evaluation must be developed that can accurately assess the complex and sophisticated skills held by effective teachers (Haertel 1991; Haney et al. 1987; Millman and Darling-Hammond 1990).

Advocates of new assessment methods argue that conventional approaches subscribe to an outdated model of teaching and learning. To such critics, underlying conventional assessment methods is an overly simplistic conception of the work of teachers. In the conventional conception, the teacher is akin to a trained technician who is responsible for implementing appropriate instructional practices that have been designed by administrators and specialists. In this view, the key objectives of teacher assessment are to insure that minimum standards concerning ability and training are met and to monitor to what extent teachers do, in fact, enact appropriate practices (Darling-Hammond 1986).

The newer thinking on teacher assessment advocates the use of a fundamentally different conception of what teaching entails and what constitutes effective teaching. In this view, effective teaching is a far more complex, specialized and broader set of processes than as conceived by conventional models and conventional assessment methods. Rather than viewing teaching as a

---

[2] Representative examples of the literature on teacher professionalization include: Carnegie Forum 1986; Corcoran, Walker and White 1988; Rosenholtz 1989; Johnson 1990; Shedd and Bacharach 1991; Darling-Hammond 1984; Weis et al. 1989; Sergiovanni and Moore 1989; Conley and Cooper 1991.

matter of implementing prescribed procedures, critics argue that teaching involves the ongoing use of judgement in the planning, conception, implementation, assessment and revision of effective teaching practices. Teachers must analyze the needs of their students, assess the resources available, take account of the goals of the school, district and parents and then devise appropriate curricular programs. The model of the teacher underlying this view is that of the highly trained, highly skilled professional (Darling-Hammond 1986).

Since the mid 1980's there have been a number of ongoing projects concerned with spelling out in more detail this newer understanding and conception of what teaching actually entails and devising alternative sets of methods by which to assess it.

One of the most important efforts to begin to develop alternative forms of teacher assessment is the work of the Joint Committee on Standards for Educational Evaluation. This group is composed of representatives from the domains of both practice and research and from a number of leading educational organizations. The latter include the American Educational Research Association, the American Psychological Association, the National Education Association, the American Federation of Teachers, the American Association of School Administrators and the National School Boards Association.

The role of this group has been to develop standards and guidelines defining what a professional model of teacher evaluation should be. Although its constituent groups have a history of strong disagreement over such issues, the committee was able to reach consensus on what constitutes sound assessment, culminating in the 1988 manual, The Personnel Evaluation Standards.

Another important initiative in this direction was the creation by the Carnegie Task Force of a National Board for Professional Teaching Standards. The task of the NBPTS is to oversee and provide leadership in the development of new definitions and methods of teacher assessment, licensure and certification. Like the Joint Committee, the Board includes practicing teachers and representatives of teachers' organizations. In 1989 the NBPTS' initial efforts culminated in a position paper, Toward High and Rigorous Standards for the Teaching Profession, designed to summarize a more complex, more specialized and more demanding view of what teachers should know and be able to do.[3]

NBPTS' Ideals of Professional Teaching

1. Commitment to Students

[3] Information about the NBPTS may be obtained from the National Board for Professional Teaching Standards, 1320 18th St. NW, Suite 401, Washington, D.C. 20036.

2. Subject Knowledge and Pedagogical Knowledge
3. Student Management and Monitoring Skills
4. Self Growth and Reflection
5. Professional Membership in School Community

The ideals elaborated in the NBPTS paper go far beyond the domains of knowledge, skills and attitudes that are typically targeted in conventional methods of teacher assessment. They emphasize the importance of professional judgement and commitment and the role of continuous learning on the part of teachers. Moreover, they suggest the depth, breadth and complexity of the knowledge and skills that effective teachers must possess.

Beyond the general guidelines and standards outlined by the Board and the Committee, a number of related projects have been concerned with elaborating, in more detail, these newer conceptualizations of what constitutes effective teaching.

Among the most prominent of these projects is the work of Lee Shulman and the Teacher Assessment Project (TAP) at Stanford University. This project has been funded by the Carnegie Foundation and has been closely linked with the NBPTS.[4]

Shulman's object has been to delineate and formalize the "knowledge base" - the particular knowledge and skills - that exemplary teaching requires. His view is that teacher knowledge is a distinct canon that, as of yet, is little understood and even less codified. Contrary to much thinking, this knowledge base is far more sophisticated and complex than the everyday wisdom common to us all by virtue of having once been a student. It "incorporates but is distinct from the general norms of civility, from the general ethics of childrearing, from the disciplines that are dealt with as subject matters in the school curriculum and from the social sciences employed in educational research and practice" (Shulman and Sykes 1986, pp.i-ii).

Shulman's Knowledge Base Categories

1. General Liberal Education
2. Subject Knowledge
3. Knowledge of Pedagogical Principles
4. Pedagogical Content Knowledge
5. Understanding of Student Diversity

---

[4] Materials from TAP may be obtained from the Teacher Assessment Project, School of Education, Stanford University, Stanford, CA 94305

6. Knowledge of Curricular Materials
7. Performance and Communication Skills
8. Knowledge of Educational Context and Foundations

The most central and distinctive category within this corpus of knowledge is what Shulman calls "pedagogical content knowledge." This is the core of what teachers must learn and use to teach well; it lies at the interface of subject knowledge, of teaching skills and of knowledge of learners. It involves knowing what works with particular kinds of students to transmit particular kinds of learning in particular kinds of settings.

Another prominent set of efforts to develop new ideas and methods of teacher assessment have come from the Center for Research on Educational Accountability and Teacher Evaluation (CREATE) at Western Michigan University, a national research and development center funded in 1990 by the Office of Research of OERI. CREATE is conducting research on existing forms of teacher assessment in an effort to develop new criteria, procedures and tools that could be used in teacher assessment.[5]

In pursuit of these efforts, CREATE, in collaboration with NCES, conducted a Fast Response Survey (FRSS) in 1993-94 to assess existing patterns of teacher assessment in public elementary schools across the country. This survey was designed to collect information on the extent and forms of prevailing methods of teacher assessment and to ascertain teachers' views of what is and should be assessed and how assessments should be used.

Among the most influential work to come out of CREATE is Michael Scriven's efforts to conceptualize the requirements of effective teaching (1991). As opposed to the highly cognitive model developed by Shulman, Scriven's "duties-based" approach is far more behavioral and performance oriented. His alternative is quite simple and also quite comprehensive. He has developed a list of the implicit duties of the teaching job, that is, what constitutes a description of what teachers are actually responsible for knowing and doing as providers of services to parents and students

Scriven's Outline of Duties of the Teachers

1. Knowledge of Subject Matter
    a. In the fields of special competence

[5] Materials from CREATE may be obtained from the Center for Research on Educational Accountability and Teacher Evaluation, Western Michigan University, Kalamazoo, MI 49008.

b. In across-the-curriculum subjects

2. Instructional Skills
    a. Communication Skills
    b. Management Skills
        i. Management of process
        ii. Management of progress
        iii. Management of emergencies
    c. Course Construction and Improvement Skills
        i. Course planning
        ii. Selection and creation of materials
        iii. Use of available resources
        iv. Evaluation of course, teaching, materials and curriculum

3. Assessment Skills
    a. Knowledge about student assessment
    b. Test construction/administration skills
    c. Grading/ranking/scoring practice
        i. Process
        ii. Results
    d. Reporting student achievement
        i. Knowledge about reporting achievement
        ii. Reporting practice

4. Professionalism
    a. Professional Ethics
    b. Professional Attitude
    c. Professional Development
    d. Service to the Profession
        i. knowledge about the profession
        ii. helping beginners and peers
        iii. work for professional organizations
        iv. research on teaching
    e. Knowledge of duties
    f. Knowledge of the school and its context

5. Other Services to the School

Scriven's model is specifically designed as an alternative to assessment tools that focus on

teaching style, such as the checklists commonly used in conventional classroom performance assessments, and also as an alternative to the more cognitive-oriented tools, such as the teacher exams.

These newer conceptualizations of teaching practice and teachers' work now being developed by researchers such as Shulman and Scriven have, in turn, created a demand for new methods of teacher assessment. As a result, there have also been a number of efforts to revise and upgrade conventional methods, such as teacher tests and use of student achievement scores, in order to take account of some of the more trenchant criticisms raised and also to take advantage of newer conceptualizations of the teaching role and knowledge base.

The leading example of the new generation of teacher examinations is the Praxis Series: Professional Assessments for Beginning Teachers currently being developed by ETS as a replacement for the NTE. Praxis is composed of three stages. The first is a computerized test of basic literacy and numeracy skills. The second is a paper and pencil test of subject matter knowledge and general pedagogical principles. The third is an observational assessment of classroom teaching performance. At this point Praxis is solely used for licensure of beginning teachers.[6]

In addition, a number of efforts are underway to address problems with the use of student performance in the evaluation of teacher performance. One of the most promising of these efforts is the work of William Sanders, a statistician at the University of Tennessee. Sanders and his associates have been working on methods of isolating the effects of schools and teachers on the achievement gains of students in the Tennessee state-wide student testing program. This approach attempts to control for the impact of student background characteristics on achievement - an ubiquitous problem - by comparing each student's deflection from their own longitudinal growth curve. Moreover, this approach also attempts to address the problem of teachers teaching to the test by the use of new tests each year (McLean and Sanders 1984; Sanders and Horn 1993).[7]

However, in addition to efforts to revise and improve conventional assessment methods, many have argued that broader, more complex, more specialized definitions of what teaching entails necessarily require new evaluation instruments that are able to capture these concepts. Much effort is being devoted to both uncovering the "true" nature of effective teaching and also finding more "authentic" means of assessing the characteristics of superior teaching. Currently, several new methodologies are being tested by a number of groups including NBPTS, CREATE and TAP.

Among the most prominent of the new methods of teacher assessment under

---

[6] Copies of the Praxis Series may be obtained from Educational Testing Service, Rosedale Rd., 02-Z, Princeton, NJ 08541.

[7] Materials on the Tennessee Value-Added Assessment System can be obtained from: William Sanders, Statistical and Computing Services, P.O. Box 1071, University of Tennessee, Knoxville, TN 37901

experimentation are the use of peer and self evaluations. The latter method, in particular, borrows from the approach to assessment commonly used in higher education. The rationale is that teachers, like other professionals, ought to police their own ranks. In one version, teachers create a portfolio, such as that used in tenure reviews at colleges and universities, which presents evidence of the teacher's accomplishments and performance. In another version of this approach, a team of peers observes a beginning teacher in the classroom in order to make promotional and other decisions.

A second method under development is the use of assessment development laboratories (ADLs) for teacher evaluation. Several prototype centers have been established by the NBPTS. The objective of the assessment labs is to use a variety of intensive evaluation exercises for the national recognition and certification of outstanding experienced teachers. In this model, senior-level teachers spend from one to three days undergoing evaluation at a center. Among the evaluation activities that could be used are lesson planning exercises, videotaped teaching performances, exercises in which teachers evaluate and critique textbooks, exercises in which teachers demonstrate the use of curriculum materials, and written examinations requiring extended essay-type answers.

These newer teacher assessment methods are currently under development or are being tested in small numbers of schools and districts. As a result, these newer methods are only beginning to be assessed. In particular, issues of validity and reliability are yet to be addressed. It is becoming clear, however, that these methods may be less amenable to standardization and, hence, more time consuming and expensive to administer than some conventional techniques. Other than acknowledgment of these kinds of concerns, there has, as of yet been little attempt to explore the strengths and weaknesses of this newer methods of assessment. The following section suggests some of the kinds of limits that these newer methods must overcome.

One of the central problems confronting assessment is how to account for the effect of the social context on teacher performance. That is, the quality and performance of teachers cannot be understood, nor evaluated, in isolation from the quality and performance of schools. Laboratory methods of assessment, such as those pioneered by the National Board for Professional Teaching Standards, are designed to clearly scrutinize specific skills and abilities of teachers. In this approach, teachers are removed from the real world of the classroom in schools and assessed in the artificial world of the lab. The strength of such experimental methods is that they allow assessors to view how well teachers perform normal activities - conceive lesson plans, use curriculum materials or present model lessons - in the absence of distractions.

But the distractions screened out of the laboratory setting may, in fact, be very pertinent factors shaping real-life teacher performance. Indeed, some teachers who perform well in the laboratory may not be able to perform well in particular classrooms. Laboratory methods of assessment do not really control, but rather ignore, the effects of social context on teacher quality. As a result, by not viewing teachers under actual classroom conditions, such methods may provide one-sided assessments of actual teacher quality. Moreover, by striving to maximize the

professional growth of outstanding teachers, such assessment methods ignore the central objective behind conventional approaches - to insure the accountability of all those in the nation's classrooms.

On the other hand, the other major example of newer methods - peer and self evaluations - are better able to account for the effects of social context. In fact, the strength of such methods is that they allow teachers to evaluate themselves in reference to standards that reflect the realities of the school context. The assumption underlying these methods is that those that actually do the job are in the best position to judge how well it could be and actually is done. The standard of comparison and, hence evaluation, is the performance of other teachers in the same or similar schools. Teachers assessed are not expected to perform any better than those that assess them - their peers.

By maximizing teacher involvement in assessment, self and peer evaluations may, however, minimize the involvement of others. It is for this reason that peer assessment methods used in higher education have been under attack in recent years. Critics have charged that universities are too research-oriented and not concerned enough with teaching or with the needs of students. One common criticism, for example, is that hiring and promotion decisions are dominated by a faculty member's research and publication performance and that teaching performance counts for little (e.g. Smith 1990). Hence, by placing evaluation in the hands of practitioners, such methods may provide one-sided assessments of actual teacher quality - favoring professional development and neglecting accountability, especially to student clients.

Given these limitations to the newer genre of methods, is the problem of teacher assessment intractable? Are the requirements of accountability methods simply not the same as those of employee development methods? Is it not possible to both hold teachers accountable and also foster their personal and professional growth? Or, are these purposes irreconcilable and mutually exclusive? There is, in fact, a large body of research rarely read or referenced by educational analysts that speaks to this very problem and debate.

A SOCIOLOGICAL APPROACH TO TEACHER ASSESSMENT

The problem of assessing teacher quality is really a subset of the larger issue of evaluation common to all organizations and workplaces. How does one fairly and accurately evaluate and assess employees or members in any setting? This issue has long been a central topic of study for those working in the field of the sociology of work and organizations.

Schools, of course, present an especially troublesome and important variant of the employee assessment problem for social scientists. Unlike the productive and technical sectors of the economy, the means and ends of teachers' work are highly ambiguous. In schools, the "production process" involves individuals working, not with raw materials or objects, but with other individuals (Bidwell 1965; Lortie 1975). Assessment is made difficult because there is no clear definition of what the final "product" is or should be and what is the best "technology" to achieve it. These dilemmas are, however, not unique to schools. Much of the service and public sectors (e.g. hospitals, municipal government, social work, etc.) face the same set of difficulties in employee and organizational assessment. In interactional work of all kinds, evaluation is particularly ambiguous. But, although the degree of difficulty and ambiguity may vary, all settings, organizations and workplaces must confront similar issues when it comes to employee evaluation and assessment (Scott 1987; Kanter 1981).

Within the field of the sociology of work and organizations, all employee and organizational assessment is inherently a normative and social activity, whether those assessed are teachers, social workers, auto plant workers, engineers or senior managers. The effort to determine what is effective performance is never value-free and, whether intended or not, involves a series of highly value-laden choices among numerous possible alternatives. Sociologists of work and organizations have insightfully delineated the range of these decisions and choices that must be confronted in employee assessment and the kinds of values and interests each choice represents. These researchers have effectively shown how different methods of assessment reflect different sets of choices concerning such categories as: the purpose of the evaluation; the domain of focus; the level of analysis; the criteria of evaluation; the type of data or information collected and used; and the viewpoint adopted. It is these different sets of choices that distinguish competing methods of assessment. These choices are not usually made explicit nor examined, but they are highly consequential. That is, most assessments are influenced substantially by sets of unquestioned premises.[1]

The list below illustrates some of the choices available to and made by the different approaches to teacher assessment. Undoubtedly there are others. The central point here is that teacher assessments are faced with a multiplicity of such options and that any one method or approach can only adopt a subset of those necessary for adequate assessment. This list can be used as a tool to establish these limits, that is to evaluate evaluations. It can be used to identify

---

[1] Representative examples of this literature include: Nord 1983. Perrow 1977; Kimberly 1983; Kanter 1981; Cameron and Whetten 1983; Goodman 1977; Zammuto 1982; Scott 1987, Pp. 319-338.

those choices any given method emphasizes and, hence, those choices each de-emphasizes. In this way, the strengths and weaknesses of a method can be elucidated.

Choices in Teacher Assessment

1. Approach
    - Accountability
    - Professional Development

2. Information Collection Technique
    - Survey Questionnaire
    - Classroom Observation Checklists
    - Teacher Examinations
    - Student Examinations
    - Teacher Interviews
    - Teacher Portfolios
    - Experiment

3. Specific Objective
    - Teacher Entry
    - Teacher Certification/Licensure
    - Teacher Promotion and Tenure
    - Teacher Termination
    - Teacher Self Development and Growth

4. Level or Unit
    - Teacher
    - School
    - District
    - State
    - Nation

5. Viewpoint
    - Self
    - Peer
    - Client
    - Administrator

6. Standard or Ideal

Skilled Craftsperson
Trained Technician
Efficient Manager
Autonomous Professional

7. Domain
   Teacher Ability
   Teacher Training/Preparation
   Teacher Performance
   Student Performance
   Teacher Professionalism
   School Working Conditions

That decisions concerning what and how to assess are both value-laden and consequential is aptly illustrated by comparing the choices adopted by those advocating greater teacher accountability versus those advocating greater teacher professionalization.

To many advocates of increased teacher accountability, school problems are, to an important extent, a result of inadequacies in the classroom performance of teachers. Teachers are held responsible and this is reflected in the kinds of assessment choices made. The target of scrutiny and ultimately, blame, is typically the ability, the training or the motivation of individual teachers. From this viewpoint, there is a need to increase the application and impact of conventional assessment methods, such as classroom observations and the use of student test gains. It logically follows that adherents of this approach look to improving schools by improving teachers, through one of any number of possible prescriptions - more rigorous entry exams, teaching workshops, remediation, merit pay or termination.

Many advocates of teacher professionalization, on the other hand, begin with a different set of assumptions. To this perspective, school problems are, to an important extent, a result of inadequacies in the school itself and the surrounding environment. In this view, focusing solely on the teacher ignores the social context within which teachers work and unfairly holds teachers responsible for problems not of their making. Inadequacies in teachers' performance may actually be symptoms of a host of other deeper causes such as, lack of time to prepare instructional lessons, mismatches between what teachers were trained to teach and what they have been assigned to teach, disruptive conditions related to problems with student misbehavior, lack of adequate teaching and classroom resources, or overly strenuous course load assignments for teachers. Adherents of this approach tend to favor assessments that are either controlled by teachers themselves (e.g. portfolios, peer observations) or that separate assessment from context (e.g. assessment labs). Finally, in contrast to the accountability approach, this alternative tends to offer a set of antidotes and prescriptions centered around improving the school and its

organization and management.

While each of these approaches to assessment shares the same overall goal - to improve education - each tends to favor different strategies, different foci, different levels of analysis and different viewpoints. It is important to identify the choices made, and hence, the choices not made, by any particular approach to assessment because these such choices make a difference. At the heart of assessments are judgments, whether implicit or explicit. These judgements are consequential; they assign responsibility and, ultimately, credit or blame.

Moreover, in truth, both approaches are probably partially correct, but neither is likely sufficient alone. Both employee accountability and employee development are important needs. The performance of individual teachers and of the schools in which they work are important. Assessments of teachers, schools, districts and states all require placement in the larger surrounding social context for comparisons to be meaningful. Finally, the viewpoints of individual teachers, faculties, and administrators are all potentially biased, but all also potentially important sources of information on how well teachers and schools work.

There is a growing consensus among sociologists of work and organizations that the goal of finding the "one best way" - the "authentic objective" measure of quality in any given setting or occupation - is misplaced. In this view, all assessment methods can potentially offer valuable information, but each one is also limited and partial. From a sociological viewpoint, the role of assessors should be to, first, make explicit the underlying, and usually implicit, choices and, second, to elucidate the strengths and weaknesses inherent in each choice. Armed with some awareness of the limits of each, the role of those charged with employee assessment should be to develop and utilize multiple measures and multiple methods to be used in conjunction with one another.

IMPLICATIONS FOR NATIONAL ASSESSMENTS

In the following section I will try to draw out the implications of the above-described approaches to the theory and methods of teacher assessment for the development of national large-scale sample surveys of teacher quality. Are any of the ideas and methods currently utilized or being developed, adaptable for use in the kinds of studies NCES? What are the implications of the current approaches and of the sociological view for national assessments? First, however, it is necessary to briefly review the purpose of national education data collection.

The role of national surveys, such as those undertaken by NCES, are distinctly different than that underlying most of the above-described activity surrounding teacher assessment. By definition, national survey assessments are less directly applied. They are not designed to either evaluate or develop the skills and performance of individual teachers or of school staffs. They are not designed to directly inform hiring, promotion or dismissal decisions. Indeed, sample surveys typically take great pains to mask the identities of individual teachers, administrators, schools and districts. Moreover, responding to survey questionnaires is ostensibly voluntary and, hence, anything suggesting control, evaluation or accountability could face problems of low response rates and high levels of respondent bias.

To be sure, underlying national survey assessments is the same general motive - to improve the quality of teaching and education - and the results of these efforts could and do speak to issues of both accountability and professional development. But, the role of large-scale surveys is in some sense even more fundamental.

In order to improve education, teaching must be improved. But, in order to improve the quality of teaching, it is necessary to first ascertain and understand how and why the teaching workforce is performing as it is. It is the production, description and explanation of these facts that is the task of large-scale surveys. By definition, surveys are equipped to provide the "big picture" either as cross-sectional "snapshots" or as longitudinal trends.

Given NCES' role, the usefulness of many of the above-described methods of assessment, especially the newer genre, is probably limited. The new professional development approach has fostered a distinct trend towards more labor-intensive, open-ended, and individualized methods of assessment, such as the use of portfolios, laboratories and peer assessment teams. These methods, of course, are not readily adaptable for use in large-scale surveys. On the other hand, as we have seen, there is abundant criticism of many of the conventional methods based on standardized, questionnaire formats (e.g. teacher examinations) - formats that are more readily adaptable for use in large-scale surveys.

NCES is, of course, capable of using observational, intensive, naturalistic, laboratory or field methodologies for data collection. But, the strength of such an agency lies in its ability to conduct large-scale, highly comprehensive, technically sophisticated, nationally representative studies. Few other organizations have the resources, capabilities and experience to produce such

data.

That this focus is a strength and not a liability is widely recognized, as the discussion at the beginning of this report attests. Nationally representative survey data are necessary to both delineate and explain overall patterns of teacher quality. Moreover, statistical analysis of data on a broad range and large number of school sites is necessary to establish appropriate and generalizable measures of levels of particular teacher characteristics and to pinpoint, with confidence, the determinants and consequences of these levels.

Thus, given the role and strengths of NCES, is the problem of large-scale survey assessment of teachers intractable? Is it simply not possible to measure something as "subjective" as effective teaching with the use of "objective" instruments?

The view that I offer here is that large-scale national survey assessments of teacher quality are quite possible. Rather than view different approaches and methods to assessment as competitors, different measures could be viewed as complementary partners.

More specifically, NCES could use survey questionnaires to collect data on a number of domains of activity in schools, from a number of different groups, representing a number of different viewpoints, and couched or analyzed at different levels. Large-scale surveys can provide valuable information on how and why the teaching workforce is performing as it is; it could seek information both on how accountable and how professionalized teacher and schools are by developing, testing, refining and analyzing a range of different measures of teacher quality. Moreover, although this information would not be collected explicitly for applied purposes, such as licensure and promotion, it could be designed to aid practitioners in the development of multiple methods, measures and instruments for more applied assessments at the local level.

The role of such NCES research could be to explore the properties, assess the limitations and study the interrelationships of a wide range of methods and measures. Rather than embark on an endless search for the one best measure, the goal could be to develop and utilize a comprehensive set of multiple measures and methods to be used in conjunction with one another.

Such a multiple-method strategy would first require an examination of a range of measures that used together would cover as many of the options suggested in the pervious list of choices as possible - the subject of the following section.

SOURCES OF SURVEY MEASURES OF TEACHER QUALITY

In this final section I will briefly summarize some examples of measures of teacher quality that have been designed for use in large-scale sample survey questionnaires. These measures can provide a useful starting point in future efforts to develop a comprehensive approach to the assessment of teachers through surveys. Together, they effectively cover many of the key domains, levels of analysis and viewpoints, discussed above. In particular, I will focus on possible measures of the actual performance of teachers, as opposed to measures of teacher ability and training. Some of these survey measures and questions are already in use by NCES or other organizations concerned with large-scale educational studies, while others are being developed by various educational research organizations.

Self Assessment

Survey questionnaires designed for teacher respondents can either assess teacher performance indirectly, by asking questions of aspects of the teachers' performance, such as practices, efforts and behavior, in order to make inferences, or directly by asking the teacher to simply evaluate themselves.

NCES' National Educational Longitudinal Study of 1988 (NELS:88) currently collects data from teachers on a number of aspects of their actual instructional practices. These include: teaching strategies and techniques, time allocation, and perceived efficacy. The teacher component of NELS:88 is not, however, nationally representative; the teacher sample is based on a national sample of 8th grade students.

Among the most promising new efforts to develop better measures of effective teaching practices for use in surveys is the Third International Mathematics and Science Study (TIMSS). As part of the background to this upcoming study, the Survey of Mathematics and Science Opportunities Project (SMSO) at Michigan State University has been developing new sets of measures of mathematics and science teaching processes and practices. This project has attempted to incorporate much of the newer thinking on what constitutes good teaching practice into the measures. The range of aspects of teaching practice for which indicators are being developed is impressive: lesson structure, textbook usage, student evaluation, materials use, teacher student interaction, homework practices, classroom management, student grouping within classrooms, teacher knowledge of curriculum policies and regulations.[9]

Besides actual practices, another set of important aspects of performance are teacher commitment and motivation. Both NCES' Schools and Staffing Survey (SASS) and The Condition of Teaching: A State by State Analysis, a survey of public school teachers conducted

---

[9] Research Reports from the SMSO (Survey of Mathematics and Science Opportunities Project) can be obtained from William Schmidt, U.S. National Research Center, TIMSS, 115 Erikson Hall, East Lansing, MI 48824-1034.

by the Carnegie Foundation for the Advancement of Teaching, both in 1988 and 1990, include questions on the degree of effort and time teachers spend on their jobs per week - possible indicators of commitment. Another highly creative measure of teacher commitment - a question asking teachers how much of their own money they spend per year on supplies to support classroom activities - was used in the 1992 Carnegie survey. Such a measure explores the extent to which teachers assume individual responsibility in the face of inadequate support.

Both the Metropolitan Life Survey of the American Teacher (MLS) and the National Survey of Science and Mathematics Education (NSSME) provide examples of direct measures of self assessment. The MLS is a series of annual national telephone surveys of public school teachers conducted by Harris Associates since 1984.[10] The NSSME is a study of science and mathematics educational practices in the U.S. conducted by the research Triangle Institute with support from the National Science Foundation in 1985 and 1993.[11]

In the 1991 MLS survey, and the 1985 NSSME teachers were asked to evaluate various aspects of their own qualifications to teach. Questions include: are you teaching courses for which you do not feel adequately qualified?; has your training prepared you to teach students from a variety of ethnic backgrounds?

Administrator and Peer Assessment

Faculty performance as a whole can be assessed by asking either a sample of teachers or an administrator from each school to evaluate overall faculty practice, behavior and effort.

The MLS surveys have collected extensive data on teachers' perceptions of school and faculty quality. These have included questions asking teachers to evaluate: the overall quality of education in the school; the dedication of fellow teachers; how well teachers are doing in the school; the quality of new teachers; the qualifications and competence of teachers in the school; the degree to which most teachers seem to care about their students; the success of the school in preparing students for education beyond high school and for jobs; the respondent's experience working with other teachers in the school

The NSSME surveys have collected data on both principals' and teachers' perceptions of faculty quality. These have included questions asking respondents to evaluate: the preparation level of the science and mathematics teachers; the articulation of instruction across grade levels; school and teacher quality; the proportion of the faculty they consider competent and those not adequately prepared to teach.

---

[10] Reports from the Metropolitan Life Survey of the American Teacher are available from Metlife: The American Teacher Survey, P.O. Box 807, Madison Square Station, New York, N.Y. 10159-0807.

[11] Materials concerned with the NSSME (National Survey of Science and Mathematics Education) may be obtained from Iris Weiss, Horizon Research, Inc., 111 Cloister Court, Suite 220, Chapel Hill, NC 27514.

The SASS asks both teachers and administrators to indicate whether teacher absenteeism is a problem at their school - a possible indicator of commitment.

Another example of assessment of faculty as a whole comes from the School Assessment Survey developed by Research for Better Schools (RBS), a federally mandated and funded educational research laboratory for the Mid-Atlantic region. In several waves throughout the 1980's, this survey gathered a wide range of information on the characteristics of public schools that voluntarily participated in the laboratory's research and consultation programs. Of interest is the inclusion of an unusually large number of questionnaire items on faculty attitudes, behavior and professionalism. Examples include, to what extent: the faculty display a sense of pride in their school; are committed to doing the best job; press students to work at higher levels; work cooperatively; maintain an interest in improving the school; try new methods in their teaching; take an interest in the social and emotional problems of their students; and do textbook teaching only.[12]

School working conditions can be assessed by asking either a sample of teachers or an administrator from each school to evaluate conditions. Numerous surveys have examples of such measures.

SASS, for instance, contains information on a host of issues concerning how professionalized both teachers and schools are. These include measures of teacher:

1. preparation - ability, education, training and background of teachers
2. recruitment - hiring requirements for teacher candidates
3. induction - entry, apprenticeship and mentoring programs for new teachers
4. utilization - assignment to service chores, and to out-of-field teaching
5. compensation - salary levels
6. resources - classizes, equipment, preparation time, aides
7. decision-making participation
8. personal and professional development opportunities

---

[12] Materials concerned with the School Assessment Survey may be obtained from Research for Better Schools, 444 N. 3rd St., Philadelphia, PA 19123.

CONCLUSION

The foregoing section has touched upon a sample of the kinds of survey questions on teacher quality and performance that currently exist. It provides a starting point in the development of a comprehensive set of survey measures of teacher quality. This report suggests that such an effort would need to take account of the multiple domains, levels, and viewpoints described in this report and brought to attention by the current approaches to teacher assessment. The next step in such an effort is to research the properties, assess the limitations and study the relations among a wide range of possible items and indicators in order to develop a set of multiple measures for use in conjunction with one another in surveys.

REFERENCES

Berk, R.A. 1988. "Fifty Reasons Why Student Achievement Gain Does Not Mean Teacher Effectiveness." Journal of Personnel Evaluation in Education. 1:345-363.

Bidwell, Charles. 1965. "The School as a Formal Organization." Pp. 973-1002 in Handbook of Organizations, edited by J. March. Chicago: Rand McNally.

Bobbitt, Sharon, P. Quinn and P. Dabbs. 1992. Filling the Gaps: An Overview of Data on Education in Grades K through 12. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Cameron, Kim and David Whetten. 1983. Organizational Effectiveness: A Comparison of Multiple Models. N.Y.: Academic Press.

Carnegie Forum on Education and the Economy. 1986. A Nation Prepared: Teachers for the 21st Century. New York: Carnegie Forum.

Carnegie Foundation for the Advancement of Teaching. 1990. The Condition of Teaching. New York: Carnegie Foundation.

Choy, S., Henke, R., Alt, M., Medrich, E., & Bobbitt, S. (1993). Schools and staffing in the US: A statistical profile, 1990-91. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Conley, S. and B. Cooper. 1991. The School as a Work Environment: Implications for Reform. Boston: Allyn and Bacon.

Corcoran, Thomas, Lisa Walker and J. Lynne White. 1988. Working in Urban Schools. Washington, D.C.: Institute for Educational Leadership.

Darling-Hammond, Linda. 1986. "A Proposal for Evaluation in the Teaching Profession." Elementary School Journal. 86:531-551.

Darling-Hammond, L. 1984. Beyond the commission reports: The coming crisis in teaching. Santa Monica, CA: Rand Corporation.

Darling-Hammond, L., and A. Wise. 1985. "Teaching Standards or Standardized Teaching?". Educational Leadership.41-2:66-69.

Dwyer, Carol and Stufflebeam, Daniel. 1992. "Evaluation for Effective Teaching."forthcoming in Handbook of Educational Psychology.

Educational Research Service. 1988. Teacher Evaluation: Practices and Procedures. Arlington, VA: ERS.

Goodman, P., Pennings, J. and associates. 1977. New Perspectives on Organizational Effectiveness. San Francisco: Jossey-Bass.

Grissmer, David and S. Kirby. 1993. "Designing the Teacher Followup Survey: Issues and Content." Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Haertel, Edward. 1991. "New Forms of Teacher Assessment." Pp. 3-29 in Review of Research in Education. Washington, D.C.: American Educational Research Association.

Haney, Walter, G. Madaus and A. Kreitzer. 1987. "Charms Talismanic: Testing Teachers for the Improvement of American Education." Pp. 169-238 in Review of Research in Education. Washington, D.C.: American Educational Research Association.

Holmes Group. 1986. Tomorrow's Teachers. East Lansing Michigan: Holmes Group.

Joint Committee on Standards for Educational Evaluation. 1988. The Personnel Evaluation Standards.

Johnson, Susan, M. 1990. Teachers at Work. New York: Basic Books.

Kanter, Rosabeth. 1981. "Organization Performance: Recent Developments in Measurement." Annual Review of Sociology 7:321-49.

Kennedy, M. 1992. "The Problem of Improving Teacher Quality While Balancing Supply and Demand." In Boe, E. & Gilford, D., (Eds)., Teacher Supply, Demand and Quality (63-126). Washington, DC: National Academy Press.

Kimberly , John. 1983. "Pondering the Performance Puzzle: Effectiveness in Interorganizational Settings." Pp.249-264 in R. Hall and R. Quinn eds. Organizational Theory and Public Policy Beverly Hills, Ca: Sage.

Kirst, M. 1990. Accountability: Implications for State and Local Policy Makers. Washington, D.C.: Office of Educational Research and Improvement, U.S. Department of Education.

Lortie, D. 1975. School Teacher. Chicago, IL: University of Chicago Press.

Madus, G. and D. Pullin. 1987. "Teacher Certification Tests: Do They Really Measure What We Need to Know?" Phi Delta Kappan. 69:31-38.

Medley, D and H. Coker. 1987. "How Valid are Principals'Judgements of Teacher Effectiveness?" Phi Delta Kappan. 69:138-140.

McLean, R. and W. Sanders. 1984. Objective Component of Teacher Evaluation - a Feasibility Study. Knoxville, TN: University of Tennessee, College of Business Administration Working Paper Series.

Millman, J. 1981. Handbook of Teacher Evaluation. Beverly Hills, CA.: Sage.

Millman, J. and L. Darling-Hammond. 1990. The New Handbook of Teacher Evaluation. Newbury Park, CA: Sage.

Milner, J.O. 1991. "Suppositional Style and Teacher Evaluation." Phi Delta Kappan. 72:464-467.

Murnane, R. & Raizen, S. (Eds.) 1988. Improving Indicators of the Quality of Science and Mathematics Education in Grades K-12. Washington, DC: National Academy Press.

Naftulin, D., J. Ware and F. Donnelly. 1973. "The Doctor Fox Lecture: A Paradigm of Educational Seduction." Journal of Medical Education. 48:630-635.

National Governor's Association. 1987. Results in Education. Washington, D.C.: National Governor's Association.

National Education Statistics Agenda Committee. 1990. A Guide to Improving the National Education Data System. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

National Board for Professional Teaching Standards. 1991. Toward High and Rigorous Standards for the Teaching Profession. Washington, D.C.

Nord, Walter. 1983. "A Political-Economic Perspective on Organizational Effectiveness." Pp. 95-130 in Organizational Effectiveness: A Comparison of Multiple Models, edited by K. Cameron and D. Whetten. N.Y.: Academic.

Office of Educational Research and Improvement (OERI). 1987. What is Happening in Teacher Testing: An Analysis of State Teacher Testing Practices. Washington, D.C.: OERI, U.S. Department of Education.

Perrow, Charles. 1977. "Three Types of Effectiveness Studies." pp. 96-105 in Goodman, P., Pennings, J. and associates. eds. New Perspectives on Organizational Effectiveness. San Francisco: Jossey-Bass.

Powell, Arthur, E. Farrar and D. Cohen. 1985. The Shopping Mall High School. Boston: Houghton-Mifflin.

Rosenholtz, S. (1989). Teacher's Workplace: The Social Organization of Schools. New York, NY: Longman.

Sanders, William and S. Horn. 1993. An Overview of the Tennessee Value-Added Assessment System. Knoxville, TN: University of Tennessee Press.

Schmidt, William. 1993. TIMSS: Concepts, Measurements and Analyses: Appendices. East Lansing, MI: Michigan State University, Survey of Mathematics and Science Opportunities Research Report Series.

Sclan, Eileen. 1994. Performance Evaluation for Experienced Teachers: an Overview of State Policies. American Association of Colleges for Teacher Education.

Scott, W. Richard. 1987. Organizations: Rational, Natural and Open Systems. Englewood Cliffs, N.J.: Prentice-Hall.

Scriven, Michael. 1991. "Duties of the Teacher" in TEMP Memos - The Teacher Evaluation Models Project. Kalamazoo, MI: Center for Research on Educational Accountability and Teacher Evaluation (CREATE).

Sergiovanni, T.J & J. Moore. (1989). Schooling for Tomorrow. Boston, MA: Allyn and Bacon.

Shavelson, R. McDonnell, L., & Oakes, J. (1989). Indicators for Monitoring Mathematics and Science Education. Santa Monica, CA: Rand Publishing.

Shedd, Joseph and Samuel Bacharach. 1991. Tangled Hierarchies. San Francisco: Jossey-Bass.

Shulman, L. 1986. "Those Who Understand: Knowledge Growth in Teaching." Educational Researcher. 15:4-14.

Shulman, L. and G. Sykes. 1986. "A National Board for Teaching? In Search of a Bold Standard". New York: Carnegie Forum on Education and the Economy.

Smith, P. 1990. Killing the Spirit. New York: Viking.

Special Study Panel on Education Indicators. 1991. Education Counts. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Stodolsky, S. 1984. "Teacher Evaluation: The Limits of Looking." Educational Researcher. 13:11-18.

Weis, L., Altbach, P., Kelly, G., Petrie, H., & Slaughter, S. (1989). Crisis in teaching Albany, NY: State University of New York Press.

Wise, Arthur E., and Darling-Hammond, L.. 1985. "Teacher Evaluation and Teacher Professionalism". Educational Leadership. 42-4:28-33.

Zammuto, R. 1982. Assessing Organizational Effectiveness.

Listing of NCES Working Papers to Date

Please contact Ruth R. Harris at (202) 219-1831
if you are interested in any of the following papers

Number	Title	Contact
94-01 (July)	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02 (July)	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03 (July)	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04 (July)	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-05 (July)	Cost-of-Education Differentials Across the States	William Fowler
94-06 (July)	Six Papers on Teachers from the 1990-91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
94-07 (Nov.)	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
95-01 (Jan.)	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02 (Jan.)	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03 (Jan.)	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95-04 (Jan.)	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
95-06 (Jan.)	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07 (Jan.)	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-08 (Feb.)	CCD Adjustment to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09 (Feb.)	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10 (Feb.)	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11 (Mar.)	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12 (Mar.)	Rural Education Data User's Guide	Samuel Peng
95-13 (Mar.)	Assessing Students with Disabilities and Limited English Proficiency	James Houser
95-14 (Mar.)	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95-15 (Apr.)	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95-16 (Apr.)	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95-17 (May)	Estimates of Expenditures for Private K-12 Schools	Stephen Broughman
95-18 (Nov.)	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96-01 (Jan.)	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-02 (Feb.)	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96-03 (Feb.)	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
96-04 (Feb.)	Census Mapping Project/School District Data Book	Tai Phan
96-05 (Feb.)	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96-06 (Mar.)	The Schools and Staffing Survey (SASS) for 1998-99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96-07 (Mar.)	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96-08 (Apr.)	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96-09 (Apr.)	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998-99 SASS	Dan Kasprzyk
96-10 (Apr.)	1998-99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96-11 (June)	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96-12 (June)	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96-13 (June)	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96-14 (June)	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96-15 (June)	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk

Listing of NCES Working Papers to Date--Continued

Number	Title	Contact
96-16 (June)	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96-17 (July)	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
96-18 (Aug.)	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
96-19 (Oct.)	Assessment and Analysis of School-Level Expenditures	William Fowler
96-20 (Oct.)	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96-21 (Oct.)	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96-22 (Oct.)	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96-23 (Oct.)	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96-24 (Oct.)	National Assessments of Teacher Quality	Dan Kasprzyk